AD-A254 699

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

...average 1 hour per response, including the time for reviewing instructions, searching existing data sources, ...g the collection of information. Send comments regarding this burden estimate or any other aspect of this ... to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson ... Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

(2)

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
01-07-92

3. REPORT TYPE AND DATES COVERED
Annual Technical 3/1/91 2/29/92

4. TITLE AND SUBTITLE Super Auditory Localization for Improved Human-Machine Interfaces

5. FUNDING NUMBERS
Grant AFOSR-90-0200
2313/CS
61102F

6. AUTHOR(S)
Nathaniel Durlach, Principal Investigator

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research Laboratory of Electronics
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER
AFOSR-TR 92 0762

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Dr. John F. Tangney
Program Manager - Life Sciences
Air Force Office of Scientific Research
Bldg. 410 Bolling AFB, Washington D.C. 20332-6448

NL

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

DTIC
S ELECTE
AUG 17, 1992
B

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

Work during the period 3/1/91 - 2/29/92 has continued on the development of a hybrid stimulation system in which a virtual auditory environment is combined with a real visual environment.

This system has been developed to help explore the effects of various transformations of auditory localization cues on both resolution and response bias. Initial research has focussed on the effects of altering the cues available to the listener for determining sound source direction in the horizontal plane. Of particular interest are alterations that magnify these cues and thus lead to supernormal performance. Although these experiments have not yet been completed, results to date indicate that current models of auditory behavior are adequate for predicting the observed changes in resolution but inadequate for predicting the observed change in response bias.

14. SUBJECT TERMS			15. NUMBER OF PAGES
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Technical Report

Work during the period 3/1/91 - 2/29/92 has proceeded along two major fronts. First, we have begun conducting adaptation experiments using a hybrid stimulation system in which the auditory environment is "virtual" (the acoustic stimuli are synthesized using a convolvotron and head tracker and presented to the listener through earphones) and the visual component is "real" (visual stimuli result form a ring of silent speakers together with lights placed in a ring around the listener at a distance of 6 ft). Second, we have continued development of the visual component of the virtual-environment (VE) system. Use of the hybrid system (as opposed to a full VE system) has allowed us to begin experimental work sooner then if we waited for completion of the VE system.

Two major classes of transformations are being employed in our experimental auditory work. In one case, the set of HRTFs remains normal but the mapping between these HRTFs and head orientation (i.e., direction of the source relative to the head) is altered. Rotations of the interaural axis constitute an important subset of this class. A second subset, which we refer to as "warping" and are now actively exploring, involves magnification of angle in the frontal sector and minification off to the sides. One such subset of warping transformations is given by the equation

$$\theta'(N,\theta) = \frac{1}{2}\tan^{-1}\left[\frac{2N\sin 2\theta}{1-N^2+(1+N^2)\cos 2\theta}\right],$$

where different transformations $\theta \rightarrow \theta'$ of azimuthal angle are achieved by using different values for N. For N = 1, $\theta' = \theta$. For N>1, the frontal region is expanded and the sides are contracted. For N<1, the sides are expanded and the front contracted (see the accompanying figure).

In the second major class of transformations, the HRTFs are themselves altered. Transformations in this class that we are now considering simulate an enlarged head. One such transformation is the exponentiation transformation discussed by Durlach and Pang (1986) in which the complex spectrum to each ear is raised to a power. A second such transformation is achieved by frequency scaling, i.e., the altered HRTF $H'(\omega,\theta,\phi)$ is defined in terms of the normal HRTF $H(\omega,\theta,\phi)$ by the equation $H'(\omega,\theta,\phi) = H(K\omega,\theta,\phi)$, where K>1.

In general, we would expect listeners to adapt more readily to the first class of transformations than the second (because adaptation to the first requires only an intersensory recalibration, not altered acoustical processing).

The experimental protocol that we are using to examine adaptation to these transformations (selected on the basis of exploratory work performed earlier in the year) consists of a sequence of interleaved training and test runs. Each test run in the sequences consists of 26 trials of a 13-alternative angle identification experiment using a click-train stimulus, azimuthal angles separated by 10 ranging from -60 to +60, and a fixed head position. No correct-answer feedback is provided in these tests so that changes in response bias (i.e., adaptation) can be measured without feedback-induced distortion. In the training runs (each of which lasts 10 minutes), the subject is required to track the supposed angular location of the sound source (which moves randomly from speaker location to speaker location and is verified by the light associated with the given speaker) by pointing his/her nose at the "active" speaker. In this manner, the subject becomes familiar with the mapping between head position and acoustic stimulus. Each session (which lasts roughly 1.5 hrs) involves the following sequence of test and training runs:

Test using normal cues (1n)
Train using normal cues
Test using normal cues (2n)
5 minute rest
Test using altered cues (1a)
Train using altered cues
Test using altered cues (2a)
Train using altered cues
Test using altered cues (3a)
Train using altered cues
Test using altered cues (4a)
5 minute rest
Train using altered cues
Test using altered cues (5a)
Train using altered cues
Test using normal cues (3n)
Train using normal cues
Test using normal cues (4n)
Train using normal cues
Test using normal cues (5n).

The initial runs with normal cues provide a control; the final runs with normal cues permit us to observe negative after-effects (as well as to ensure that our subjects leave the experiments with normal hearing -- a requirement of the human-use committee). Taken all together, this sequence provides us with 5 tests using normal cuesand 5 using altered cues in each session.

After initial exploratory work with this experimental procedure, formal tests were conducted with one warping transformation (the above equation with N = 3) and one exponentiation transformation (squaring of the complex spectra), using 4 subjects for each transformation. We have completed roughly 8 sessions per subject for the first transformation and 2 sessions per subject for the second transformation. We are now completing these experimental sessions and attempting to analyze and interpret the results. Although these experiments have not yet been completed, results to date indicate that current models of auditory behavior are adequate for predicting the observed changes in resolution, but inadequate for predicting the observed changes in response bias.

Work on the development of the visual component of the VE system has included 1) experimenting with graphics software tools with the aim of selecting the best library for real time interactive graphics development, 2)evaluating different head-mounted displays (HMDs) and head trackers, 3)trying to find a means of generating NTSC video output from the DECstation 5000 with which to drive the HMD, and 4)considering various forms of interconnection between the graphics computer and the audio simulation computer.

Software development has proceeded in parallel at a number of different levels. At the lowest level of programming, a graduate student has built up a 3-D graphics package in C using the X library. It provides an intuitive point-and-click user interface for construction of rudimentary virtual worlds. Moving parts have not yet been implemented, but easily can be. This low-level approach provides the control necessary to minimize delays and flicker by updating only the part of the scene that has changed, and using double buffering. However, the code is extremely complex, making modification difficult, and it is impossible to utilize the 3-D geometry accelerator and scan converter of the DECstation when making only Xlib calls. For this reason, the shading algorithms that have ben implemented in this package are too slow for real-time use, and only wireframe graphics are a possibility when just using Xlib.

At the highest level, a demonstration package using PHIGS is being developed by a programmer who has recently joined our staff. While PHIGS has many advantages and enables shaded surfaces to be generated as quickly as wireframe drawings, it is impossible to update a scene without rerendering everything, and double buffering is not available. Therefore, image motion tends to involve both flicker and unacceptably long delays.

At this time, both programmers are beginning to investigate PEX, which seems to offer most of the advantages of both the low-level and high-level approaches. An initial demonstration has been written showing fast animation with shaded polyhedra. Problems with reading head-tracker data in this programming context, as well as generating stereoscopic views, are currently being addressed.

Getting binocular images from the graphics engine to the HMD remains a problem. No reasonably priced RGB-to-NTSC converter has been found with the capability to perform area-of-interest conversion on a window corresponding to one eye's view. Experiments using CCD cameras aimed at windows on the screen have found it difficult to get good images due to the scan rate mismatch. However, there is still some possibility of synchronizing the camera, monitor, and HMD. There is also a possibility that DEC will soon be making a video converter board, or that a Truevision scan converter can be suitable modified.

During the coming year we intend not only to explore further transformations, but also variations in experimental procedures. In one such procedure, we hope to estimate the amount of adaptation to a magnified interaural axis using a method of adjustment in which the observer rotates his head and adjusts the magnification factor until the source remains stationary during head rotation. Other variations include changes in motor system involvement on response selection and communication (e.g., head pointing vs hand pointing vs verbal or keyboard responding) and alterations in the amount and type of visual information provided to the listener.

Of particular importance for future work is the inclusion of a head-tracking system with a sufficiently large work space to permit translation as well as rotation. The mechanical head tracker we began to develop during the past year for this purpose has not proved to be adequately reliable to be used without further work. It is also worth noting that the Bird we purchased from Ascension for head tracking purposes has evidenced a design bug previously unknown to the designers. Thus, most of our current work is being conducted with a Polhemus (shared with our associates at BU).

We are also concerned about the extent to which our results are limited by the actions of the Convolvotron (specifically, the method used for interpolating angle in the Convolvotron). Apart from cooperating in a study with our associates at BU in an analysis of the Convolvotron, we hope to develop a number of other systems for achieving appropriate signals. Systems now under consideration include a much simplified synthesis system and a system in which the alterations are achieved by modifying the head of a slave robot in a teleoperator system.

Concerning the development of the visual component of our system, the staff programmer will focus on writing fast virtual environment software in PEX, and the graduate student will pursue the integration of the whole system, modifying or building whatever hardware is necessary to accomplish the generation of the stereo NTSC images, and developing software to enable a 386 to function as the central coordination center of the system by processing data from any of several head-trackers and dispatching it over ethernet connects to the graphics engine, to the computer housing the Convolvotron, and possibly to a slave robot for teleoperator applications.

With regards to the communication of our ideas and results during the past year, note the following:

(1) "Auditory Localization in Teleoperator and Virtual Environment Systems: Ideas, Issues, and Problems", by N. Durlach, has been published in Perception (Vol 20, Number 4, pp 543-554, 1991)

(2) "On Externalization of Auditory Images", co-authored by N. Durlach and a whole gaggle of associates, has been accepted for publication in PRESENCE.

(3) Two invited talks on material associated with this grant were presented during the past six months:

"Sensing and Displaying Acoustic Information", N. Durlach, ILP Symposium on Telerobotics, MIT Oct 29 - 30, 1991.

"Super Auditory Localization for Improved Human-Machine Interfaces", N. Durlach, DOD User-Computer Interaction Technical Group, Nov 5, 1991 San Antonio, Texas.

(4) A further invited talk, "Super Auditory Localization Displays", by Durlach, Held, and Shinn-Cunningham, has been prepared for a session on Auditory and Tactile Displays, International Symposium Seminar and Exhibition, Society for Information Display, May 17 - 22, 1992, Boston, Mass.

(5) A paper on preliminary experimental results will be presented at the ASA meeting in New Orleans by Ms. Shinn-Cunningham.

(6) Work connected with this grant is also influencing thinking at other agencies via Durlach's participation in meetings with individuals at ONR (e.g., Hawkins, Allard) and NASA (e.g., Jenkins).

Finally, it should be noted that our work during the past year was unexpectedly slowed by the resignation of one of our staff, Xiao Dong Pang. Due to severe financial pressures associated at least in part with the addition of 3 new members to his family (2 parents and 1 child), Dr. Pang left the engineering field and accepted a much higher paying job with a financial "trading" company in New Jersey. Although Dr. Pang formally took a leave of absence from MIT, we do not expect him to return. Furthermore, we do not plan to replace him unless further funds are obtained for this (and associated) research. In part because of this loss of personnel and in part because we believe that the importance of personalized HRTFs has been overplayed (see our forthcoming note on "Externalization of Acoustic Images"), we did not during the past year spend time and money on the measurement of such HRTFs.

Nose faces 0 degree, clockwise is positive

Y(n,x) = atan(2n[sin(2x)]/[1-n*n+(1+n*n)cos(2x)])/2

Y(1,x) = x, Y(n,-x) = -Y(n,x), Y(n,0) = 0, Y(n,90) = 90

Y'(n,0) = n > 0, Y'(n,90) = 1/n, Y'(n,x) >= 0

Y''(n,x) has sign of 1-n